

BLM LIBRARY



88055087

June 1985

EMS16536.41R

CALIFORNIA DESERT AIR WORKING GROUP

DATA BASE STUDY

VOLUME 3 OF 7

DATA SELECTION

ENVIRONMENTAL MONITORING & SERVICES, INC.*

Formerly Environmental Monitoring & Services Center

of Rockwell International Corporation

2421 West Hillcrest Drive

Newbury Park, California 91320

(805) 485-8771

TD
883.5
.C2
E68
1985
v.3

RAY OF COMBUSTION ENGINEERING, INC.

Table of Contents

	Page
1.0 Data Selection Criteria	1
2.0 Selected Data	3
3.0 Statistical Analyses.	4

BLM LIBRARY
BLDG 50, ST-150A
DENVER FEDERAL CENTER
P.O. BOX 25047
DENVER, COLORADO 80225

CALIFORNIA DESERT AIR WORKING GROUP (CDAWG) DATA BASE STUDY:
DATA SELECTION

In the course of this program, air quality, visibility, and meteorological data have been received from various sources and of differing quality, completeness, and format. The compilation of a computer accessible data base in a common format will require converting the data format as well as the transcription (keypunch, digitizing, etc) of data in hardcopy form to machine readable form. Thus, it will not be cost efficient to include data of questionable value in the data base. Further, screening the data before inclusion in the data base will result in a more compact data base.

1.0 Data Selection Criteria

The selection of data to be included in the data base is based upon a quantitative procedure which uses a ranking of the data quality, comparability, and representativeness in addition to the effort required to compile the data. By necessity, however, the actual data selection will be somewhat subjective. Our philosophy is to err on the conservative side and include data so long as undue effort is not required for the task. Since the data quality, representativeness, comparability, etc. have already been rated by the comprehensive indexing scheme described in the Task 1 report, the end user of the data base will have little difficulty in selecting data according to any specific overall data usefulness criteria. Thus, our approach in selecting data for inclusion in the data base will be whether the effort required is reasonable for the overall value of the data to be included. With this in mind, the data will be selected according to the following criteria; data not meeting the criteria will be excluded.

Data on Computer Tape. All data on computer tape satisfying the following minimum conditions:

* At least one year's data

- * Processed (as opposed to raw) data

- * For air quality and visibility data, information must be available to estimate the data quality and comparability of the data, even if the quality index ends up being poor

Since most of the available data sets include multiple sites, even if one site satisfies the above criteria all the data will be included because the effort required to read and convert data on computer tape is the same whether part or the whole tape is read. The advantage of this approach will be that the maximum possible data will be available and the deletion of those unsuitable will be up to the end user.

Hardcopy Data (Strip charts, printout, etc.). Transcribing hardcopy data on to computer tapes will require substantial effort. Thus, we will be more selective in including hardcopy data. In addition to the criteria listed above for tape data, hardcopy data must satisfy the following criteria:

- * Must be spatially and temporally unique data, i.e., data for the same parameter and data for the same geographic area and same time period are not available on tape. From the end user point of view, if data are available on tape for a particular geographic area and time period, the additional effort required to transcribe a nearly duplicate hardcopy set data will not be cost effective.

- * Must be air quality data. The thrust of this program is towards an air quality data base, and any needed visibility and meteorological data will be available on tape. Thus, the few

hardcopy meterological data will not be transcribed.

Unlike multi-site computer data, for obvious reasons, only the data for specific sites satisfying the above criteria will be transcribed onto computer tapes.

Special Study Data. These will be included only if the variables are those commonly used for air quality and meteorological data reporting and if the measurement methods are also routine rather than esoteric research methods.

Unusual Data Format. The criteria for including data in unusual data formats will be the same as those for hardcopy data.

Emissions Data. The latest emissions inventory available for the California desert is dated 1979. This, revised to the extent possible, will be included in the data base and submitted separately.

2.0 Selected Data

On the basis of the above data selection criteria, the following preliminary list of data will be included in the data base:

1. Data from all California Air Resources Board air monitoring sites in the SE Desert and the Great Basin Valleys Air Basins;
2. All air quality and meteorological data from all non proprietary Southern California Edison air monitoring sites in the California desert;
3. All CO data from all CALTRANS sites in the Imperial Valley;

4. Meteorolglcal data from the Weather service for Bishop, China Lake, Cuddeback Lake, Edwards Air Force Base, George Air Force Base, Needles, Blythe, Silver Lake, Thermal, and Mojave
5. All air quality, visibility, and meteorological data from Naval Weapons Center (NWC) sites.
6. All air quality, meteorological, and visibility data from Army sites- Ft. Irwin and Death Valley
7. All meteorological and visibility data from EPA Las Vegas (EPA-LV) sites- Ft. Irwin, Iron Mountain, and Edwards Air Force Base;
8. All meteorological data from San Diego Gas and Elelctric sites; and
9. Meteorological data for Lancaster from SCAQMD

Some of the data for the NWC, Army and EPA-LV sites may be duplicates or data for the same site for different time periods. Duplicate data will be weeded out during data compilation. With the exception of the Weather Service data, data from all availble periods will be included. Only post 1960 Weather Service data will be included.

3.0 Statistical Analyses

After the data base has been compiled, statistical analyses will be performed. Although exhaustive analyses are not planned, the statistical analyses will meet the following objectives:

1. Characterization of data base, i.e., means and frequency distributions;

2. Spatial distribution of data;
3. Temporal behavior of data; and
4. Correlations

To meet the above objectives, the following statistical analyses are proposed:

1. Calculation of monthly, seasonal, and annual means and standard deviations for O_3 , SO_2 , oxides of nitrogen, CO, and TSP at the following sites:
 - Imperial;
 - Indio;
 - Barstow;
 - Victorville;
 - Trona;
 - Lancaster; and
 - Bishop (TSP only).
2. Calculation of monthly, seasonal, and annual mean and standard deviation of visual range, contrast or B- scat at all visibility monitoring sites.
3. Normal and log normal frequency distribution plots of the parameters and sites listed in 1 and 2 above.
4. Isopleth maps of the entire desert for O_3 , SO_2 , oxides of nitrogen, CO, TSP, and vsisbility using annual averages in 1977 and 1981.
5. Monthly time series plots of the parameters and sites listed in part 1 above for 1977 and 1978; wind directions will also be plotted where available.

6. Monthly time series plots of visibility data at all visibility monitoring sites.

7. Visibility roses:

Visibility vs. wind direction

Visibility vs. relative humidity;

this will be done for all sites which measure visibility and meteorological parameters concurrently.

In our original proposal we had also planned to do factor analysis to determine the contributions of various sources or source categories. However, after examining the data base, we have decided that factor analysis would not be as cost-effective as the statistical analyses listed above, and therefore it will not be performed. We had originally expected to find a more complete data set than we found, but in fact there are few sites at which multiple variables are measured concurrently for extended periods. Therefore, the correlations which we had planned for factor analysis were not feasible. Furthermore, we had originally planned to concentrate upon the relationships between visibility and particulate chemical composition, but we found that collocated visibility and particulate measurements are rare in the California desert.

BLM LIBRARY
BLDG 50, ST-150A
DENVER FEDERAL CENTER
P.O. BOX 25047
DENVER, COLORADO 80225



